DOCUMENT RESUME

ED 122 583 FL 006 453

AUTHOR Robinson, Jane J.
TITLE Performance Grammars.
SPONS AGENCY Advanced Research Projects Agency (DOD), Washington, D.C.
PUB DATE 75
CONTRACT DAHC04-72-C-0009
NOTE 26p.
AVAILABLE FROM Academic Press, Inc., 111 Fifth Avenue, New York, New York 10003 (HC $17.00). Available as part of the publication entitled "Speech Recognition."

EDRS PRICE MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS *Grammar; Language Research; *Language Universals; Language Usage; Language Variation; *Linguistic Competence; *Linguistic Performance; *Linguistic Theory; Speech

ABSTRACT

Evidence is offered to support the view that linguistic competence cannot in principle be divorced from linguistic performance in order to abstract universal properties of grammars, that rules of grammar inevitably incorporate perceptual strategies and constraints, and that grammaticality and acceptability are related to predictability. A theory of systematic variation affords direction for gathering data on rule-governed language use and for representing the results in formal grammars that predict speech behavior. Some of the strategies and constraints operating in performance and the rule-governed regularities they produce are demonstrated in the analysis of seven tape-recorded task-oriented dialogs. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ERIC

ED122583

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL BY MICROFICHE ONLY HAS BEEN GRANTED BY
Academic Press, Inc.
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

To appear in:

The invited papers of the IEEE Symposium on Speech Recognition, held at Carnegie-Mellon University April 15-19, 1974.

PERFORMANCE GRAMMARS⊛

by
Jane J. Robinson
Stanford Research Institute
Menlo Park, California

0. INTRODUCTION

There is a passage in Through the Looking Glass where Alice is being tested by the White Queen and the Red Queen to see if she is worthy to become a queen herself. The White Queen asks her: "What's one and one and one and one and one and one and one and one and one and one?" Alice says, "I don't know. I lost count." And then the Red Queen says, "She can't do Addition." The Red Queen was obviously wrong to infer from a single performance that Alice was incompetent to 'do Addition'.

It was Chomsky who elevated the distinction between what one is able to do and what one actually does on a given occasion to the status of a fundamental distinction in linguistic theory. In explicating the notion of 'grammatical rule' (Chomsky 1961), he proposed to use the term competence to refer to a speaker's implicit and intrinsic knowledge of his language and the term performance to refer to the use he makes of that knowledge at particular times. He also proposed to divorce the notion of grammar from all concern with performance; grammar was to be a reflection solely of competence. His stance in this respect goes far beyond the customary acknowledgment of the need to normalize linguistic data before attempting to state the systematic regularities that are discernible in it. His claim is that it is not possible to understand performance without first abstracting the universal properties of the grammars of natural languages from the flux of occasional utterances and thereby gaining insight into the innate structure of the human mind.

A more recent view is that universal linguistic properties are really only manifestations of broader cognitive principles that underlie the observed regularities in behavior. (Kuno 1973; Kimball 1973) In this latter view, the internalized rules of our implicit grammars are not innately given but have evolved from the interplay of cognitive processes with our experiences in communicating with each other. As a result of the interplay, the rules incorporate perceptual strategies and constraints; therefore it is impossible to eliminate performance factors from

FL006453

⊛The work reported herein was sponsored by the Advanced Research Projects Agency of the Department of Defense under Contract DAHC04-72-C-0009 with the U.S. Army Research Office.

grammars. Some of the evidence for this view is presented in Section 1 (Rules and strategies) in the course of a more detailed examination of Chomsky's initial statements concerning the division between competence and performance. The next section (Section 2. Rule-governed variation) exhibits an extension of the notion 'rule of grammar', proposed by Labov, that permits explicit incorporation of data from performance. In Section 3 (Context-sensitive performances) several rule-governed regularities discernible in seven recorded dialogs are analyzed to expose the strategies and constraints operating in actual performances.

1. RULES AND STRATEGIES

Although he is ultimately interested in the human mind and its intuitions about language, the connection Chomsky asserts between grammars and the mind is subtle and indirect. A grammar of a language describes the intrinsic competence of an ideal speaker-hearer in the sense that it generates the sentences he is theoretically capable of producing and comprehending. It is a theory of sentences, not of utterances. Since grammar is timeless and performance takes time, his grammar does not describe even the ideal speaker's performance. The White Queen asks, "What's one and one and one..." and stops at ten, but her grammar allows her to go on droning "and one and one and...." If she is tireless and immortal, she may never finish. Then all the other sentences she could have uttered will go unperformed.

There is at least one dismaying flaw in the concept of grammar divorced from performance. As long as grammarians are encouraged to sweep awkward phenomena into a dustbin labelled 'performance' and send it off to the psycholinguists, their grammars will generate sentences that no speaker ever utters, no hearer readily understands, and almost every language user rejects as ungrammatical unless coaxed by a linguist. A classic example appears in (1). Given a little time and a little linguistic sophistication, people can comprehend it and paraphrase its content, as in (2). Few would judge (1) to be acceptable, but the claim is that there is no way to exclude it from the grammar of English except on grounds of difficulty of performance, and such grounds are not admissible in grammars that are wholly concerned with competence.

(1) the house the cheese the rat the cat the dog chased caught ate lay in was built by Jack

(2) the house in which lay the cheese that was eaten by the rat that was caught by the cat that the dog chased was built by Jack

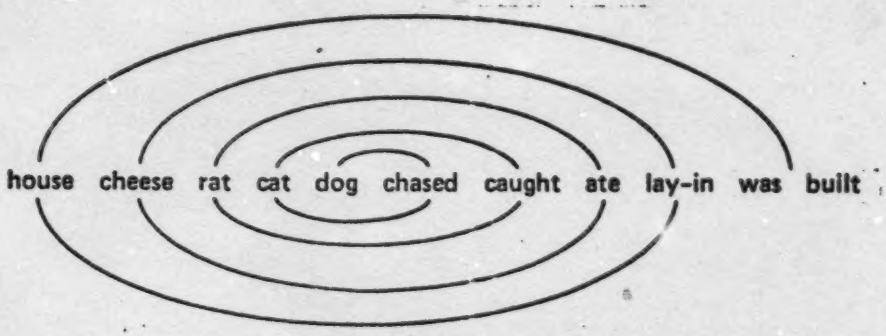
Chomsky argues that this strange state of affairs is "no stranger than the fact that someone who has learned the rules of multiplication perfectly (perhaps without being able to state them) may be unable to calculate 3,872 x 18,694 in his head, although the rules that he has mastered uniquely determine the answer." (Chomsky 1961, pp.7-8) The analogy is not very

compelling. One cannot give an example of a product generated by rules of multiplication that is both correct and 'unacceptable'.

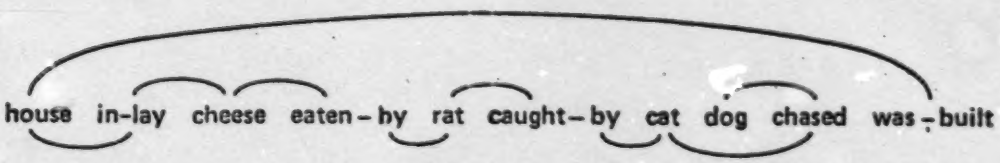
Perhaps a better analogy is that of chess. For each piece, a rule in the grammar of chess defines its legitimate moves. But there are also strategies that guide the choices of a competent player, and there is one general requirement that partakes of the nature of rule and strategy. It is the meta-rule that says, "No matter what the other rules say about how this piece can move, they cannot be applied if the result is to put your king in check." An approximate counterpart in the language game is: "Don't use rule combinations that create incomprehensible sentences."

Looking again at sentences (1) and (2), we see that their propositional content is the same; they are close paraphrases. In the difficult sentence, the structure is deeply self-embedded. In the easier one, optional transformations have inverted the relative clauses to produce perceptually simpler right-branching constructions. Some of the surface structure differences between the two are represented in (3), which shows how the embeddings of (2) appear to have been flattened in figure 3b.

(3)



(a)



(b)

A line drawn above the sentence connects a head noun with the verb it serves as subject; a line beneath the sentence connects it with the verb or preposition for which it is object.

As a result of flattening, sentence (2) is easy to comprehend. It can be parsed on the run without ove taxing one's short-term memory. (In a recent article, Kimball (1973) claims that performance limitations on short-term memory make it impossible to parse the constituents of more than two 'open' S's at the same time. Sentence (1) requires parsing five at once.)

In addition to being flatter, (2) consistently marks the beginnings of constituent boundaries with function words that signal the type of constituent likely to follow. Lacking such cues to the structure of (1), a hearer is in danger of being led down the garden path of a false parse as he processes it, so that he misconstrues the initial series of noun phrases as a single conjoined noun phrase subject. Disillusionment comes when a verb is encountered where a conjunction was anticipated, and he finds that re-parsing is necessary if the sentence is to be properly understood. This consequence, following from the deletion of function words, shows that deletion transformations, like flattening transformations, are related inseparably to performance as well as to competence. To give another example, deletion of the function word that when it introduces an embedded complement clause in subject position is unacceptable and likely to be judged ungrammatical because the result regularly misleads the hearer into processing the embedded clause as if it were the main clause of the sentence. (See the examples in (4) below). On the other hand, it is acceptable and grammatical to delete a that when it introduces an embedded clause in object position, where the class of the preceding predicate predicts the possibility of embedding.

(4) a. *he was angry frightened everyone
b. everyone was afraid (that) he was angry
c. that he was angry frightened everyone
d. it frightened everyone that he was angry

A study of how people process sentences like these has led Bever (1970, p.20) to propose that at least some grammatical rules inevitably obey behavioral constraints because "children will tend not to learn rules which produce speech forms that are hard to understand or hard to say. Thus, whatever aspects of cognition are utilized directly in speech perception will be reflected in certain properties of linguistic grammars."

It is a truism that we tend to see what we are prepared to see and hear what we are prepared to hear. This truism translates into a general linguistic principle that acceptability and grammaticality are related to predictability. The principle shows up in the statistics for texts that have not been 'made up' to illustrate linguistic points. Consider again the preceding example: that he was angry frightened everyone. It is grammatical and the first word signals the possibility of encountering an initial embedded clause. However, the possibility is even more strongly signalled if the embedded clause is extraposed to follow the predicate frightened as in (4d), it frightened everyone that he was angry. It is noticeably rare to find embedded clauses in initial subject position, as if speakers and writers intuitively avoid introducing complex sentences too abruptly with too little advance warning. In examining seven extended dialogs in casual speech (to be described more fully in a following section), I failed to find even one

embedded clause in subject position, although there were many embedded object and predicate complement clauses. The examples in (5) are representative; each comes from a different dialog. Places where a that was deleted are marked with a #; retained but deletable thats are enclosed in parentheses. The predicate that precedes and predicts the embedding is underlined.

(5) a. Your first comment was (that) the diagram didn't match ...
b. The purpose for those washers is (that) the motor has to be able to slide ....
c. Before you start, be sure # you turn each of those off.
d. Well, that means # we'll have to remove ....
e. I don't suppose # the consultant knows where ....
f. Make sure # the groove in the flywheel lines up with ....
g. Is it correct (that) the strap is attached to the pump?

If the concept of grammar divorced from performance fails us as a theory of sentences, it is completely irrelevant as a theory of sequences of sentences in a discourse. The timeless competence grammar stops at the boundary of each unuttered sentence that it generates and starts over. It can enumerate sentences in some canonical order, but cannot generate them in a functionally appropriate order, even thought the syntactic structure, word choices, use of anaphora, and selection of prosodies of every sentence are affected by past and anticipated utterances. Judgments of the acceptability of sequences of sentences are clearly related to predictability, even though we cannot predict precisely in all cases. For example, if from the sentences in (6), all ordered pairs of (a) followed by (b) are formed, we can predict which pairs will be judged acceptabie and which will not. To follow what did John do? with Bill was hit by John is predictably inappropriate. So is the sequence I know Tom hit Harry, but who hit Bill?: John HIT him (with high stress on hit). Some may hesitate to apply the term 'ungrammatical' to sequences of grammatical sentences, and yet it is difficult to say in what respect the reaction to such sequences differs from the reaction to the single sentence: I know Tom hit Harry, but John HIT Bill.

(6) a. What happened?
What happened to Bill?
What did John do?
I know Tom hit Harry, but who hit Bill?
I know Tom hit Harry, but what did John do?

b. John hit BILL.
JOHN hit Bill.
JOHN hit him.
John HIT him.
BILL was hit by John.
Bill was hit by JOHN.

If, in the examples of (6b), the pronoun him is equated with Bill, then the propositional content is the same for all of the examples. Their surface differences arise from options provided by grammar for paraphrasing the propositional content in order to foreground some portions and background others. These options serve a strategic function in the use of language, allowing us to deploy the resources of the language so that complex meanings can be delivered gradually and partially, without loss of coherence. We can take into account, as the information is being delivered, what has already been said--what is 'old' information, and link the new information to it. We can do this because the relationship between the old and the new is signalled by the choices among systematic grammatical options available to speakers for organizing their sentences. (Cf., for example, Halliday 1970; Sgall et al. 1973.) The reason why Bill was hit by John sounds odd following what did John do? is that the use of the passive in this case puts the item John into the foreground, highlighting it as if it were newly introduced into the discourse, while at the same time placing Bill in the background as if it represented old, and therefore predictable information. When it follows What happened to Bill?, on the other hand, the same sentence is acceptable and the use of the passive enhances the coherence of the two sentences in sequence.

2. RULE-GOVERNED STRATEGIES

We have seen that the passive transformational rule functions in two similar roles, making sentences more comprehensible by flattening the structure of a single complex sentence and by restructuring the content of an independent sentence to relate it coherently to the content of preceding sentences. The passive transformation is optional; the possibility of applying it or not applying it introduces variation into the structure of the same propositional content. Although it is optional, we can predict to some extent when it is likely to be applied.

The notion of systematic rule-governed variation has recently emerged as the central theme of a new paradigm for linguistic research. (Bailey 1972; Bailey and Shuy 1973.) An important part of the new paradigm is Labov's extension of Chomsky's notion rule of grammar' to include the notion of 'systematic varition' (Labov 1969). It is not possible to give an adequate treatment of the theory and methodology of 'variable rule' grammars here. I will instead give one example to illustrate the concept and the conventions for writing variable rules, to show how it is possible to retain the explicitness of generative rules while accommodating the variability of observed performance. (For extended treatments, see Labov, 1972, 1973; G. Sankoff, 1972; Cedergren and D. Sankoff, in press.)

The example is a kind of 'fast speech' rule that asserts that /t/ and /d/ are optionally deleted before a word boundary. The rule is clearly too general. While the final /t/ in the phrase kept going, might be deleted in fairly slow speech the final /t/ of got away is likely to persist at much higher speeds. The contextual features that favor deletion or retention of /t/ and /d/ are given by a variable rule:

(7) $-t,d \rightarrow \phi \Big/ \left\langle \begin{matrix} +cons \\ +cont \end{matrix} \right\rangle \left\langle \begin{matrix} \phi \\ + \end{matrix} \right\rangle \text{—\#\#} \left\langle \begin{matrix} +cons \\ -voc \end{matrix} \right\rangle$

Variability is indicated by angle brackets. Angle brackets enclosing the rewritten element on the right of the arrow means "is variably rewritten as." Combined with 0, this means "is variably deleted." Angle brackets in the contextual part of the rule enclose lists of features or categories whose presence affects the application of the rule, favorably or unfavorably. Elements not enclosed in angle brackets are obligatory.

The rule states that the most favorable environment for deletion of a word-final dental obstruent is in a monomorphemic form in which it is preceded by a continuant and followed by a word beginning with a consonant, as in hold back. The least favorable is the environment in which it functions as a past tense morpheme and precedes a word beginning with a vowel, as in missed it. There are various intermediate likelihoods of deletion, exemplified in (8).

(8)

a.	hold back --> hol back	hold off --> hol off
b.	last man --> las man	last one ---> las one
c.	kept going --> kep going	kept all of it ---> kep all of it
d.	missed me --> miss me	missed it --->miss it

From empirically established frequency counts of /t/, /d/ deletion in various environments, a statistical method derives probability coefficients for each contextual element independently. Given an input string, the values for the contextual elements actually present in the string are inserted into the formula, which combines them and calculates the probability that the dental obstruent will be deleted. Cedergren and D. Sankoff (in press) view the probabilities associated with the rules as "properly part of competence" and claim that performance is "a statistical reflection of competence."

The example given in (7) for a variable rule contains only linguistic variables in the contextual part. However, the method extends to any contextual factors, linguistic or extra-linguistic, assumed to affect frequency of rule application systematically. Covariation of application frequency with differences in age, sex, class, and style of speech have been studied for rules in various languages by various investigators. (See references previously cited.) It is not necessary for the contextual factor to appear in the body of the rule as a feature or category; its probability value can appear in the formula for predicting application, once the relevant contexts have been identified and frequencies have been established. In principle we could incorporate any systematic linguistic behavior into a grammar, even including slips of the tongue and hesitation pauses, both of which have been shown to be rule-governed (Fromkin 1971; Goldman-Eisler 1973). More practically, we now have a theoretical frame to direct the gathering of data from real performances on the frequency of application of well-established transformational rules like the passive, so that we can study the contexts which appear to favor

or disfavor their application. It is a frame that allows us to include the extra-sentential discourse contexts of what was said before and after.

In the next section we will look at some actual performances, noting the contexts that appear to shape the utterances, but before we do, I would like to illustrate more specifically how the passive transformation, revamped as a variable rule, might operate to produce the flattening of the kind of self-embedded structures we looked at earlier. The relevant context for rule application is given in the SD (structural description) part of (9). SC gives the structural change. Note that the presence of a V preceding the V affected by the transformation is a favoring context for applying the rule. The examples below the rule show some effects of application and non-application. It is assumed that the rule applies cyclically from deepest to least embedded sentence.

(9) Variable Passive Transformation

SD: NP1 X (V1) (V2) ... (Vi) Vj NP2 X

SC: NP2 B E Vj -EN BY NP1 X (V1) (V2) .. (Vi) X

a. (i) the dog chased(V) some cat
(ii) some cat was chased(V) by the dog

b. (i) the cat the dog chased(V) caught(V) some rat
(ii) the cat chased(V) by the dog caught(V) some rat

c. (i) the rat the cat the dog chased(V) caught(V) ate(V) some cheese
(ii) the rat the cat chased(V) by the dog caught(V) ate(V) some cheese
(iii) some cheese was eaten(V) by the rat the cat chased(V) by the dog caught(V)

If we start the cycle with input (9a.i) and Passive is not applied, we enter the next cycle with input (9b.i), which favors application because it has a V immediately before the affected V. If Passive is still not applied, we enter the next cycle with input (9c.i), where the presence of two Vs before the affected V makes application still more probable. At some point in very deeply nested structures, the combined effects of preceding Vs should predict application with a probability of one, indicating that no competent speaker has been observed to fail to apply the transformation in that context.

3. CONTEXT-SENSITIVE PERFORMANCES

I propose now to look at some concrete linguistic performances to see what kinds of syntactic constructions occur, what functions they perform, and how they co-vary with features of the context. These are the data needed for generative performance grammars that assign probabilities when given a proposed utterance and the contextual features in which it is embedded. As Halliday has pointed out, if we know the context, it is surprising how many features of the language turn out to be relatable to it. "... not ... that we know what the participants are going to say; [but] ... we can make sensible and informed guesses about certain aspects of what they might say, with a reasonable probability of being right. There is always, in language, the freedom to act untypically but that in itself confirms rather than denies the reality of the concept of what is typical." (Halliday, in press.)

The performances to be examined are the seven dialogs previously cited. They belong to the central paradigm for communication: the dialog that occurs when two people are working together at a common task whose nature motivates them to communicate in order to get it done. In each of the seven dialogs there is an expert and an apprentice, and the apprentice is doing some work under the partial guidance of the expert. Conditions differ with respect to vision and channel linkage (see Table 1). In two dialogs the expert can see what the apprentice is doing at any moment; in one he cannot see at all, and the apprentice must describe objects and states of affairs; and in one, vision is limited. In all four of these dialogs, the channel is speech. In another situation, the participants do not communicate directly, but through a monitor, who relays messages between them. The expert uses writing as the channel for communication, while the apprentice uses speech. The monitor accepts messages through either channel and converts them to the other. This interposes a time delay. In this situation the expert cannot see what the apprentice is doing unless he requests that a television camera be turned on and directed to a specific area. Three dialogs were collected under these conditions.

TABLE 1. Conditions for the Seven Dialogs

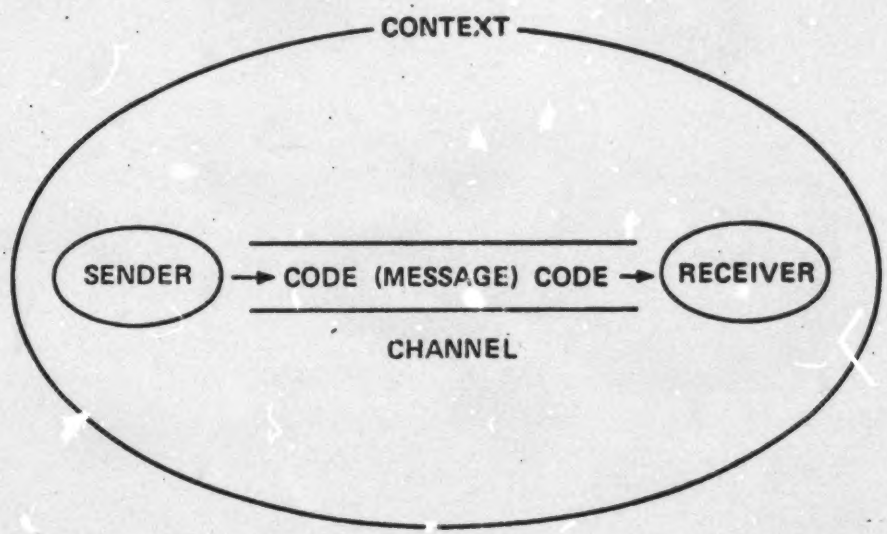
Dialog	Vision	Monitor	Expert	Apprentice
1	yes	no	RF	JT
2	limited	no	RF	FW
3	yes	no	RT	RF
4	no	no	RT	WP
5	on request	yes	JK	DN
6	on request	yes	JK	PB
7	on request	yes	RF	JP

The style or 'register' of all the dialogs except the last is that of casual speech. None of the speakers is self-conscious. Once they become task-oriented, there is no discernible awareness that someone is observing their linguistic behavior. The situation appears to solve what Labov calls the Observer's Paradox: "To obtain the data most important for linguistic theory, we have to observe how people speak when they are not being observed." (1972).

Since the dialogs were not controlled, generalized observations based on them are necessarily impressionistic. We are making some trial runs, to expose the overall problems with a view to designing more controlled experiments later to collect more tractable data. We are interested in seeing the full range of interaction between two participants when the main external structuring is imposed by the nature of the task, and then seeing what changes occur under the distancing effects imposed by lack of shared vision and by the monitoring delays. One of our research goals is the design of a performance grammar for man-machine dialog. We would like the grammar to generate a subset of English that is 'habitable' for the user. That means that we need to consider how the total system hangs together for a given type of discourse and how to excise part of it without severely disrupting the rest.

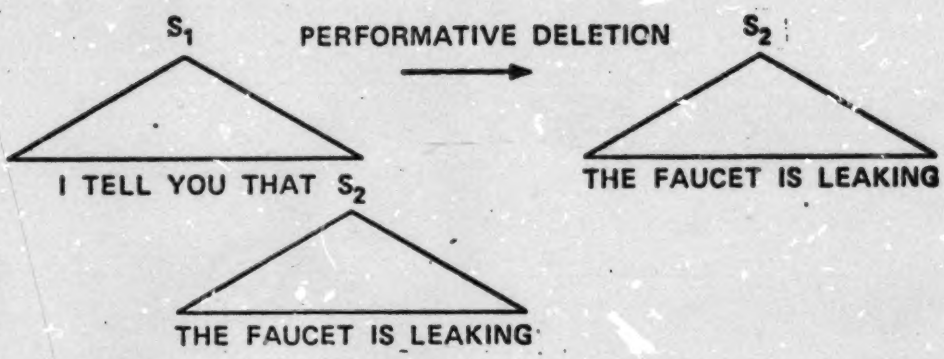
Our examination of the dialogs will be guided by the model shown schematically in Figure 2. A linguistic performance requires at least two performers, at least one message, in a code, transmitted through a channel, all of them embedded in a context. In the schematic representation for a single message, I have distinguished the two performers as 'sender' and 'receiver', rather than as 'speaker' and 'hearer', because the channel may be writing as well as speech. Some models of communication distinguish other elements as well, but this one offers a convenient way of looking at a linguistic performance with a view to getting data for a performance grammar. It is an adaptation of the model of Buhler (1934) as modified by Jakobson (1960). Their schema represents context as if it were an element distinct from the rest. I think it is important to emphasize that the context embeds the other elements. The relationship between sender and receiver is part of the context that affects the shape of the message: the message changes the context even during the course of its delivery, becoming part of the context for messages to come. It also changes the relationship between sender and receiver. This is a dynamic model, and a performance grammar is sensitive to all the features of the changing context.

(10) A COMMUNICATION MODEL



Although all the elements shown in (10) are 'there' in every single linguistic performance, one or more may be overtly emphasized, and the shape of the message reveals this emphasis in syntax, lexical choice, and prosody. For example, exclamations emphesize the sender. They give more information about his internal state than about his relation to the receiver or to the external context. Interrogatives and imperatives generally emphasize the receiver. Declaratives are the most unmarked form. According to the performative analysis, every underlying structure of a declarative is of the form I tell you (that S), where a highest S embeds the S that becomes the sentence. (Ross 1970.) Elements of the highest S are deleted transformationally, eliminating overt signs of sender and receiver from the utterance, unless the forms I and you show up in the lower S. (See (11)) This is a 'deep structure' acknowledgment of the pragmatic fact that sentences are to be uttered and heard.

(11)



The forms I and you are not pronouns like it, they, he, she, etc.; they are indexical expressions denoting the sender and receiver of a message during its transmission. When they occur in declarative sentences they are often mixed in varying degrees of emphasis with references to the external context. Compare, for instance, the opening utterances from two different dialogs. ('A' preceding an utterance identifies its sender as an apprentice.)

(12) a. A: I have this faucet here and water seems to be coming out at the wrong place ... water is leaking out from that flat horizontal surface and I don't think it's supposed to .... So can you help me fix that? (Dialog 3).

b. A: The faucet's leaking around the base of the spout. (Dialog 4).

In (12a) the sender reports his internal state; that is, his reactions to the external context along with information about that part of the context. He ends with an explicit appeal to the receiver coded in an interrogative sentence. In the second example there is no overt I and the appeal to the sender is not explicit. The interpretation of the declarative statement as an appeal is influenced by the apprentice/expert relationship of sender to receiver.

There were a surprising number of exclamations in the dialogs, but expressions of internal states of participants sometimes took the form of an almost impersonal report of relevant data. The range is from the classic example of a four-letter oath to relieve the feelings of an apprentice who brought his hand down hard on some exposed nails through expressions of less painful surprise and on to reports of worry entirely relevant to the task orientation of the context.

(13) a. A: Rats, I made a mistake. (Dialog 2).

b. A: I found a small screw on the floor which always makes me nervous when I work on machines but I'm almost certain it's not from the compressor.

E: Show me the screw, please. (Dialog 5)

In the second example we have another case in which a declarative is interpreted as an appeal for advice because of the relationship of the sender to the receiver.

A typical series of messages in the dialogs are exchanges in which the expert directs or advises the apprentice to do something and the apprentice acknowledges receipt of the message. Okay serves the latter function most frequently. The apprentice then initiates a new exchange by indicating that the directive has been fulfilled. Here, again, okay fills this function frequently, but whether with a consistently different intonation from the one accompanying its utterance in the other function is not yet clear. It is a possibility worth investigating. Embedded in these exchanges may be a subordinate exchange in which the apprentice asks a question about a current directive or points to a problem encountered in trying to fulfill it, and the expert replies. Deutsch (1974) analyzes the discourse structure of the dialogs, relating it to the structure of a 'workstation' task model. I shall concern myself mainly with the syntax.

Because of the nature of the task and the relationship between the participants, many messages are oriented towards asking for and giving information, especially the kind of information called 'advice.' When paired messages having request/response functions with this orientation are analyzed, it becomes clear that they are often a single unit syntactically and semantically as well as functionally. For example, the syntactic relationship of each request/response pair in (14) is so regular that it is possible to specify a simple series of transformations that takes the pair as input, outputting a single well-formed declarative that is an answer to the request. The elliptical responses occupy syntactic slots and contain the semantic features defined by the request, with occasional redundant overlap with parts of the request.

(14) a. A: Which side do you call the front?
E: The side with the Sears label on the tank. (Dialog 7)

b. A: Which tools should I use to get the bolts that are hard to unscrew?
E: Use the 1/2" box wrench and the 1/2" combination wrench. (Dialog 6)

c. A: How tightly should I install this pipe elbow ....
E: Only snugly. (Dialog 7)

d. E: What are you doing now?
A: Using the pliers to get the nuts in underneath the platform. (Dialog 7)

e. A: Should I take it all the way off?
E: Take it all the way off. (Dialog 3)

For (14a) the transformations are particularly simple. The request is converted to declarative form by undoing the interrogative inversion and eliminating the auxiliary DO; the you becomes I with the exchange of speakers, and the response replaces the WH-marked NP of the request:

I call WH-side the front →
I call the side with the Sears label on the tank the front.

In (14a) the word side occurs in both members of the pair: in the WH-marked NP of the request and in its replacement NP in the response. In (14b) the replacement for which tools does not contain an occurrence of the word tools, but does contain nouns referring to objects for which the word tool is a superordinate term. This is a general pattern with respect to replacements involving other parts-of-speech as well. In (14c) the WH-marked phrase is an adverb of degree -- how tightly, which is superordinate to only snugly. In (14d) the pro-verb and the what signal that the request is to be satisfied by specifying an activity. The conversion is: I AUX doing what → I am using the pliers. The pro-verb DO is superordinate to the use that that replaces it.

The request in (14e) has no overt WH element. Requests like this are sometimes called Yes/No or 'polar' questions. However, there may be more than two alternatives, so that a more general form underlying the type is the WH-OR or whether interrogative. In the example, a 'yes' would have been appropriate, but a simple 'no' leaves unresolved whether the apprentice should leave it on or take it off or take it part-way off or take some other action. Consequently it is appropriate for the response to repeat the words of the request as it does; it could even have been prefaced with the full you should. In fact, it would not be wrong in any of the examples to spell out the full declarative form of the answer instead of giving the information elliptically, but to do so would be redundant precisely because the request contains the syntactic and semantic specification for everything in the response except for what is different in the response, the new information.

Speaking generally, one concludes from the examples that a performance grammar must include a theory of pairs of sentences that constitute an exchange between participants in a speech event, in order to account on the one hand for the well-formedness of elliptical utterances that when judged by a competence grammar alone are ill-formed and incomplete, and on the other hand for their predictability

Another grammatical regularity that is discernible in the dialogs is the regular set of relationships holding among certain declaratives, interrogatives, and imperatives, whenever the declaratives and interrogatives contain the indexical words I and you, a performative verb denoting the conveying of information, and the auxiliary should or an equivalent, have to.

The close relation functionally between interrogatives and imperatives is easily noted. Interrogatives like "would you show me the pressure register so I can identify it?" (Dialog 7) have been cited as 'polite' imperatives. But the dialogs show a relation that both functionally and syntactically goes beyond an adjustment that merely makes an order palatable. Where the function to be performed is satisfaction of a need for information, the request may just as easily take the syntactic form of either an interrogative or an imperative: what are you doing? or tell me what you are doing. Every interrogative can be converted to an imperative with the addition of tell me just as every imperative can be converted to an interrogative with the addition of would you, with appropriate adjustments in word order in each case. By an equally simple transformation, every imperative can be converted to a declarative with the addition of you should, you have to, I would suggest that you, or I would suggest V-ing. See, for example, "I would suggest holding the washer and nut in one hand..." (Dialog 7).

These transformational relationships allow requests and responses to show up in various syntactic guises, depending on optional deletion of variable amounts of material. They are illustrated in (15).

(15) a. I tell you that you should tell me {what I should do / whether I should do X} (REQUEST)

→ you should tell me {what I should do / whether I should do X} (DECLARATIVE)

→ tell me {what I should do / whether I should do X} (IMPERATIVE)

→ {what should I do / ϕ should I do X} ? (INTERROGATIVE)

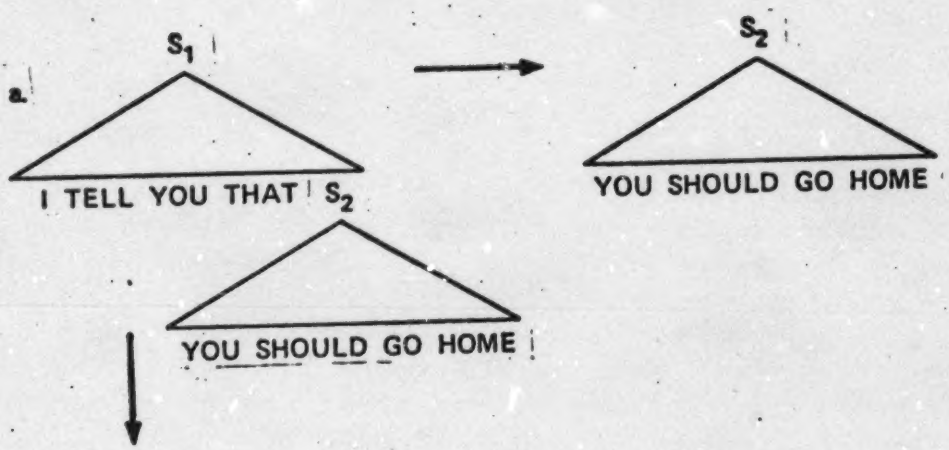
b. I tell you that you should turn it off (RESPONSE)

→ you should turn it off (DECLARATIVE)

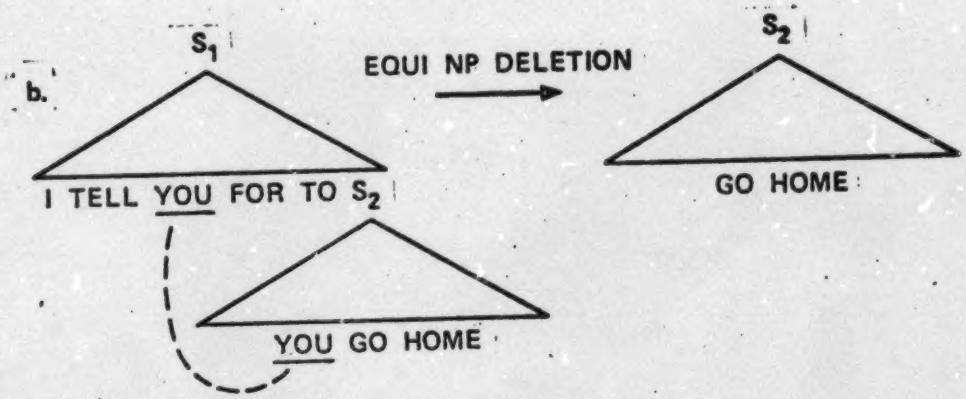
→ turn it off (IMPERATIVE)

Those who have followed the field will recognize, in (16) below, that the transformations involved are independently well motivated. That is, the transformation that produces the imperative from an underlying I tell you for you to go home is like the one producing you should go home from I tell you that you should go home, where the embedded sentence is introduced by the that-complementizer rather than by the for-to-complementizer. In this case, the declarative is generated by deleting the performative elements of the highest S, exactly as proposed by the performative analysis for all declaratives. The imperative involves, in addition, the well-known Equi-NP deletion transformation, which deletes coreferential NPs in structures like that of

(16) a.



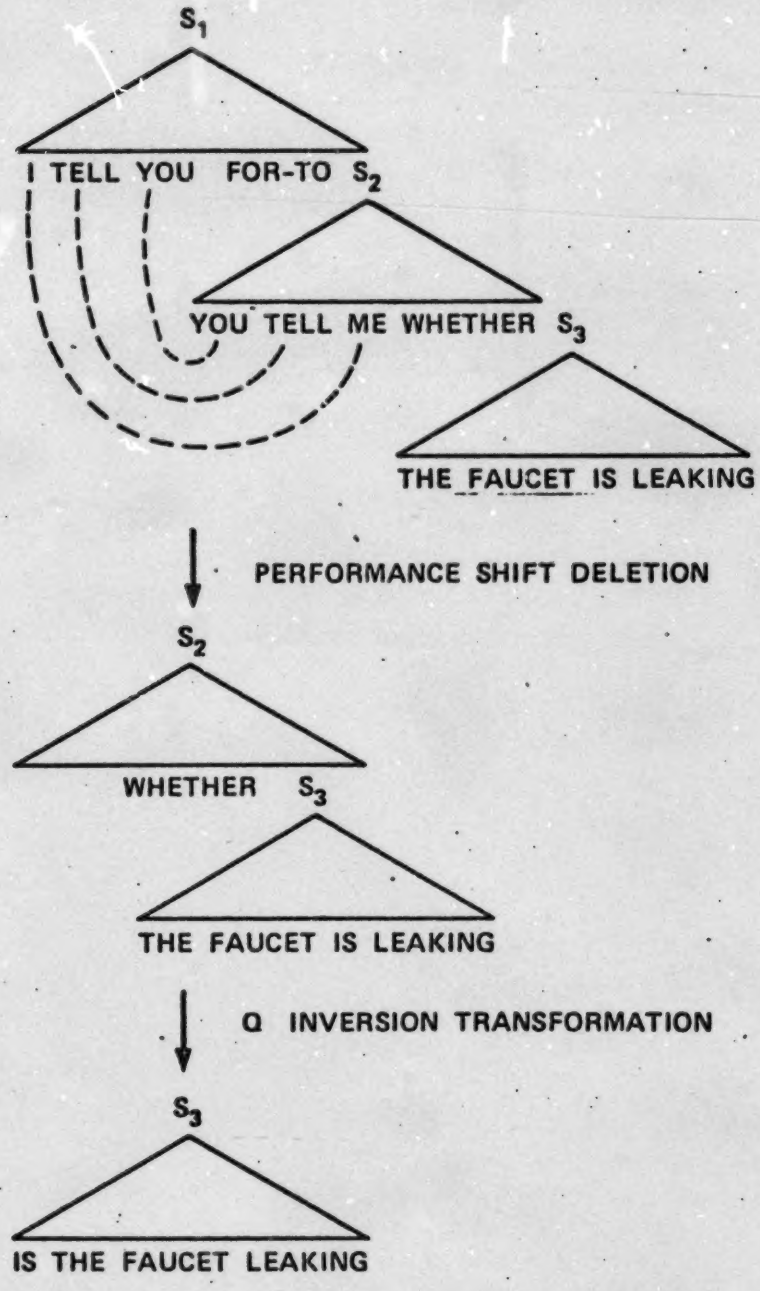
b.



The seemingly more complicated transformation shown in (17) involves the same principles--deletion of performative and coreferential material. The highest S1 (the performative S) and the S2 embedded in it contain the same predicate, tell and the same indexical I and you; but the roles of the I and you are shifted in S2. I have called the transformation 'Performative Shift' because its effect is to shift the roles of speaker and hearer.

The main point is that the syntactic type of the matrix sentence does not always distinguish between a request, a response to a request, or a neutral statement. If the message is a request, that function may be signalled syntactically by the presence of a WH-marked form in a subordinate clause. However, previous examples have illustrated the influence of extra-linguistic facts, e.g., the apprentice/expert relationship between sender/receiver as a contextual feature that favors the interpretation of a message as a request. A good example of the combination of a request signal in a subordinate clause and a covert appeal for assitance shows up in the following exchange:

(18) A: I don't suppose the consultant [the expert] knows where I put the wrenches.
E: Didn't you put them back on the table?
A: As a matter of fact I did. (Dialog 5)


S1
I TELL YOU FOR-TO S2
YOU TELL ME WHETHER S3
THE FAUCET IS LEAKING
PERFORMANCE SHIFT DELETION
S2
WHETHER S3
THE FAUCET IS LEAKING
Q INVERSION TRANSFORMATION
S3
IS THE FAUCET LEAKING

(17)

Notice that the response is in the form of an interrogative. Part of the interpretation of the first utterance as a request depends on other kinds of extra-linguistic knowledge as well: the knowledge of the importance of keeping track of tools when engaged in a task.

Whether-type questions depend on inversion for marking as interrogative. However, in English as in other language, it is possible in speech to signal this kind of question by intonation: it comes off? rather than does it come off. There are many examples of this kind of marking when apprentice and expert were in voice contact. Usually there were other marks or clues as well, as in:

(19) E: And these we turn off ... you turn to your right.
A: You mean counterclockwise?
E: It would be clockwise.
A: Clockwise. Okay. (Dialog 4)

In this example you mean is a clue, as well as intonation. It is unusual for a sender to tell a receiver what the receiver means; it is usual to ask.

The exchange in (19) illustrates some other functions toward which messages are oriented markedly. One is code orientation. In (19) the apprentice is checking the meaning of the phrase 'to your right' by offering a recoding. This is corrected and he echoes the correct word. This echoing of parts of the message previously sent was quite common when apprentice and expert spoke directly to each other but were not in eye contact.

When the expert could not see what the apprentice was doing, the means for code checking were of course affected. Much of our common code for concrete objects is established by pointing or deixis. When this means is withdrawn, verbal description becomes necessary. A striking example shows up in the comparison of two dialogs, one where the expert could see and the other where he could not.

(20) a. E: ... you have a top piece with a knurled section that you can take ahold of ...
A: What's a knurled section?
E: You've got your fingers on it. (Dialog 3)

b. E: Now underneath that is what they call a cap assembly. It has a knurled face around it right above the spout itself.
A: What does 'knurled' mean?
E: Little lines running up and down on it so you can take ahold of it. (Dialog 4)

Another example of the effects of lack of vision appears in (21), where it is combined with effects of lack of a shared coding convention. The crucial phrase is the deictic the front. 'In front' of an object is multiply ambiguous. It may mean 'between the object and sender or receiver' when either of them is facing it, or it may mean 'forward of the front' of the object itself.

messages put in buffers were delivered out of sequence. On the next session, a strategy of requiring the apprentice to address the monitor and wait for a 'yes' before proceeding proved unworkable. On the last dialog, the monitor simply interrupted the apprentice when a message was underway from the expert, but the problem was not entirely solved and some mixups still occurred.

It appears as if on some occasions the simultaneous sending of messages occurred because both participants got nervous if nothing was coming across the channel, and both were set to go off after an interval of silence that was roughly the same for each. Long silences also made the monitors nervous. Both monitors adopted the strategy of starting to relay the expert's longer messages before they were complete, and to signal by intonation that there was more to come even though the pauses were abnormally long. Unfortunately, we cannot study the acoustic detail of these interestingly deviant intonations for the light they could shed on the function of intonation in signalling sustension or finality. Since these dialogs were trial runs, we had not tried to tape them under good sound conditions. However, the phenomenon is probably reproducible, since different monitors produced it independently.

Dialog 7 records the time of transmission of each utterace. The reader might try the effect of placing a sixteen second delay between the announcement, "The mounting bolts are tight as is all the plumbing," and the reinforcing phatic response, ... "good." Or a twenty-four second delay between, "How tightly should I install this pipe elbow that fits into the pump?" and the reply ....." only snugly."

5. THE SEVENTH DIALOG

We had expected that the variations in style and syntax among the dialogs would be influenced more by the differences in the conditions under which they were conducted than by the individual differences of the participants. Impressionistically, this is true. The first two dialogs have some common features that set them slightly apart from the others, and this is probably attributable to two factors: The three different participants knew each other well and were all professionally interested in artificial intelligence. Many of their utterances were introspections about what was going through their minds as they solved the problem in front of them. However, it was relatively easy to extract the utterances that were relevant to the immediate task and these were much like the ones in Dialogs 3 and 4 where the apprentices and the expert had never previously met and where the expert was by profession a plumber.

For six of the dialogs, some obvious differences were attributable to the variables of vision and monitoring. When I first saw the transcript of the seventh dialog, the only one at which I had not been present, I was struck by the formality of the style and redundancy of the syntax. The exchanges in (18) exemplify these features.

The front of an object is established by convention. The front of a car may be behind a sender or receiver; it is still that part of the car that gets there first when the car is travelling normally along its major horizontal axis. The front of a book is what the reader gets to first when he reads it in normal order, and which part of the book that is, in relation to the book's covers, depends on whether the book is written in English or Hebrew. In the exchange of (21), the apprentice cannot interpret the phrase 'in front' because he does not know the expert's convention for orienting the compressor.

(21) E: I assume you positioned the pump so that the longer protrusion, the oil drain, is pointing toward the front.
A: Which side to do you call the front?
E: The side with the Sears label on the tank. (Dialog 7)

One of the functions of utterances that was sometimes overtly marked is keeping the channel between sender-receiver open and ready to use, even when information content is low. Such utterances are said to have a phatic function. They include greetings and politeness forms and other social and ritual expressions. They also include the encouraging murmur of the receiver, often overlapping the incoming message, so that the receiver is simultaneously receiving and sending.

Many of the utterances in the dialogs appear to emphasize this function; for example, "yeah, yeah," "okay," "right," "fine." Lack of vision alone did not affect the incidence or type strongly, but when direct voice contact was lost and there was a delay in response caused by insertion of a monitor, the effects were noticeable. There were expressions of anxiety about the functioning of the channel: e.g., "Can you hear me?" addressed by the apprentice to the monitor, who was waiting for a long response from the expert to be typed out before relaying it. (Dialog 5). It was quite apparent to experts, apprentices, and monitors that the okays and all rights had an important function. As a matter of fact they have at least two distinct functions. One is to signal that the channel is working successfully and the message has been received; the other is to signal readiness to receive a new message.

The utterance most overtly marked for successful reception was "gotcha" (Dialog 3). This function sometimes combined with the previously noted function of checking the decoding by partial repetition of the message just sent. Sometimes the repetition had an intonation and pronunciation that seemed imitative of the previous message and not characteristic of the current sender. This may also have been 'recognition by synthesis,' but it sounded more like checking and reassuring the other sender-receiver.

The second function, signalling readiness to receive a new message, proved unexpectedly important. There have been several studies recently (e.g., Yngve, 1970) of the visual and verbal cues by which people recognize whose turn it is to talk. In those dialogs where direct contact was broken between the two participants, there was considerable confusion when both tried to send messages at the same time. In Dialog 5, the first of the indirect monitored dialogs, some

(22) a. E: Are you now attaching the tubing to the elbow?
A: The elbow and tubing installation is completed.

b. A: Is it correct that the strap is attached to the pump by one of the cylinder head bolts?
E. Yes, that is correct.

Listening to the tape confirmed the impression of formality; the speech was more deliberate, the intonation less varied than that of other dialogs.

I wondered if all the differences were due to the fact that the apprentice was more knowledgeable about compressors than the others and did not know the expert. Finally someone told me of a third variable: This apprentice thought the expert was a computer. When they told him the expert was another human, he responded with an appropriate exclamation. The opening exchanges of Dialog 7 are appended.

Appendix

Protocol text of session beginning Wed. 20 Feb. 1974, 10:38

10:39:13 E: Good morning. I would like for you to reassemble the compressor.

10:39:41 A: All right I'm supposed to just start doing that and ask any questions?

10:40:01 E: Yes, I may ask you what you are doing as we go along. I suggest you begin by attaching the pump to the platform.

10:40:40 A: Seems reasonable.

10:41:49 E: What are you doing now?

10:41:58 A: Using the pliers to get the nuts in underneath the pump platform.

10:42:30 E: Ok. I assume you positioned the pump so that the longest protrusion, the oil drain, is pointing toward the front.

10:43:17 A: Which side do you call front?

10:43:29 E: The side with the Sears label on the tank.

10:43:45 A: All right.

10:43:58 A: Do you have an alternate better way to get the nuts in underneath the platform?

10:44:16 E: I would suggest holding the washer and nut in one hand and the screw in the other hand. Start the nut that way.

10:47:34 E: I realize this is a difficult step. Have you any of the nuts started?

10:48:04 A: I'm tightening the bolts now. They are all in place.

10:48:20 E: Good.

10:49:37 A: How tightly should I install this pipe elbow that fits into the pump?

10:50:00 E: Only snugly. The elbow must be connected to the tubing that connects to the tank. The end of the elbow that attaches to the tubing must be pointing toward the tubing. That direction should determine the tightness.

10:51:20 A: Fine.

10:51:42 E: Are you now attaching the tubing to the elbow?

10:51:58 A: The elbow and tubing installation is completed.

10:52:17 E: Check the nut where the tubing is attached to the tank. It may need to be tightened.

10:52:49 A: Yes I tightened that.

10:52:58 E: Good.

10:53:36 A: I'm left with a small rectangular plate that it isn't obvious what to do with.

10:53:36 E: The small metal strap is a brace for the pump and should be attached to the belt housing and the top of the pump.

10:53:57 E: That's the strap, or brace.

10:54:20 A: Yes.

10:54:51 A: Is it correct that the strap is attached to the pump by one of the cylinder head bolts?

10:55:09 E: Yes, that is correct.

10:56:37 E: This step completes work on the front side of the compressor.

10:56:57 A: Yes I've finished installing the strap now.

REFERENCES

Bailey, C-J N., Comment on papers presented at the first plenary session. In Sociolinguistics: Current Trends and Prospects. 23rd Annual Georgetown Round Table. Monograph Series on Language and Linguistics, Number 25. Edited by R. Shuy. Georgetown University Press, Washington, D.C., 1973, pp. 89-98.

Bailey, D-J. N., and Shuy, R. (Eds.), New Ways of Analyzing Variation in English, Georgetown University Press, Washington, D.C., 1973.

Bever, T.G., The influence of speech performance on linguistic structure. In Advances in Psycholinguistics. Edited by G. B. Flores d'Arcais and W.J.M. Levelt. North-Holland Publishing Company, Amsterdam-London, 1970, pp. 4-30.

Buhler, K., Sprachtheorie, Jena, 1934.

Cedergren, H.J. and Sankoff, D., Variable rules: performance as a statistical reflection of competence. Language, in press.

Chomsky, N., On the notion 'rule of grammar.' In Structure of Language and Its Mathematical Aspects. Proceedings of Symposia in Applied Mathematics, Volume XII. Edited by R. Jakobson. American Mathematical Society, 1961, pp. 6-24.

Chomsky, N., Aspects of the Theory of Syntax. M.I.T. Press, Cambridge, Massachusetts, 1965.

Deutsch, B.G., The structure of task-oriented dialogs. Proceedings of IEEE Symposium on Speech Recognition, Carnegie-Mellon University, 15-19 April, 1974.

Fromkin, V.A., The non-anomalous nature of anomalous utterances. Language, 1971, 47, pp. 27-52.

Goldman-Eisler, F., Pauses, clauses, sentences. Language and Speech, 1972, 15, pp. 103-113.

Halliday, M.A.K., Language and social man. In Papers of the Programme in Linguistics and English Teaching, Second Series, Longmans, London (in press).

Jakobson, R., Linguistics and poetics. In Style in Language. Edited by T. Sebeok, M.I.T. Press, Cambridge, Massachusetts, 1960, pp. 350-377.

Kimball, J., Seven principles of surface structure parsing in natural language. Cognition, 2(1), pp. 15-47

Labov, W., Contraction, deletion, and inherent variability of the English copula. Language, 1969, 45, pp. 715-762.

Labov, W., Some principles of linguistic methodology. Language and Society, 1972, 1, pp. 97-120.

Labov, W., Where do grammars stop? In Sociolinguistics: Current Trends and Prospects. 23rd Annual Georgetown Round Table. Monograph Series on Languages and Linguistics, Number 25. Edited by R. Shuy. Georgetown University Press, Washington, D.C., 1973, pp. 43-88.

Ross, J. R., On declarative sentences. In Jacobs, R. and Rosenbaum, P. (eds.), Readings in English Transformational Grammar. Waltham, Mass: Blaisdell, 1970, pp. 222-272.

Sankoff, G., A quantitative paradigm for the study of communicative competence. Preprint from the Conference on the Ethnography of Speaking, Austin, Texas, 1972.

Yngve, V., On getting a word in edgewise. Papers from the Sixth Regional Meeting, Chicago Linguistic Society, 1970, pp. 567-577.